30th Annual INCOSE international symposium
Cape Town, South Africa
July 18 - 23, 2020

Applying Systems Thinking to Frame and Explore a Test System for Product Verification; a Case Study in Large Defence Projects

Rune André Haugen
Kongsberg Defence and Aerospace AS
P.O. Box 1000 3601 Kongsberg, Norway
+47 93 43 14 53
rune.andre.haugen@kongsberg.com

Mo Mansouri
University of South-Eastern Norway
P.O. Box 235 3603 Kongsberg, Norway
+47 35 95 29 82
mo.mansouri@usn.no

Abstract. The test system is a vital part of delivering a verified product to the end customer. The test system used in Kongsberg Defence and Aerospace (KDA) to test missile products today needs to change to be able to cope with future requirements for faster project execution and running more projects simultaneously. This article uses a Systems Thinking approach to see the bigger picture and to ensure understanding of the entire problem domain. The system consists of the following structural elements: Data Preparation System, Mission Planning System, Simulators, Data Analysis System and Storage System. The stakeholders of the test system are testers, system owners, project managers, company, customers, government and suppliers. Several possible value added processes are foreseen to make this necessary transition; automation of test execution and test analysis to avoid bottlenecks, verification on both core product and adaption product level for modularity, combining test arena input over different systems/sub-systems/components for re-use of data, and Machine Learning only to trigger necessary manual analysis. These changes will influence the system in several ways and levels, which a possible implementation need to consider. Regarding aspects like facilities, environment, security and safety not to cause issues for the changes in question is important. The main steps in the current test process is that the test system will provide scenario data to the tester to run test scenario to generate test results to the analyzer to perform test results analysis to achieve a verified product. The test structure is a limiting factor in the process of ensuring test maturity. The analysis structure is a limiting factor in reaching the desired verification level. The test structure and analysis structure are leverage points of the test system, which can significantly change the test system. The test system should have an automated test execution and test results analysis process, not requiring tedious manual operations. The automated test process should further introduce Machine Learning to change the focus of everything to managing the exceptions. KDA will increase its probability of success in future projects by applying the proposed changes to its test system.

Introduction

Background

KDA has run three main missile projects over the last decades, which is the Penguin, the Naval Strike Missile (NSM) and the Joint Strike Missile (JSM). The Penguin development project started out in 1961, and continued through several increments to 1990. The NSM development project started in 1996, and lasted 13 years. The JSM development project started in 2008, and finish is set to 2021. The later missile development projects build on the previous, but have evolved to be quite different. A lesson learned from these three large projects is that the actual products should have been modular with core and adaption functionality to be able to re-use relevant parts in future projects. This has not been necessary in these three projects since they have been well funded and not overlapping, but that does not seem to be the case for future projects. The NSM Block 2 development project is now in its start-up phase, already pulling on resources still required in the JSM project.

A product must be subject to verification testing for the customer to accept compliance to stated requirements. The supplier and the customer must agree upon the verification testing. A verification plan states what to test, while a test description gives the details of how to execute the tests. The test system creates test reports after test execution and analysis, to document compliance with requirements.

A definition of verification is "*the evaluation of whether or not a product, service, or system complies with a regulation, requirement, specification, or imposed condition*" (Project Management Institute 2018).

The verification steps are part of the system engineering Vee Model process before operation and maintenance, as shown in Figure 1 (Osborne et al. 2005, p. 20).

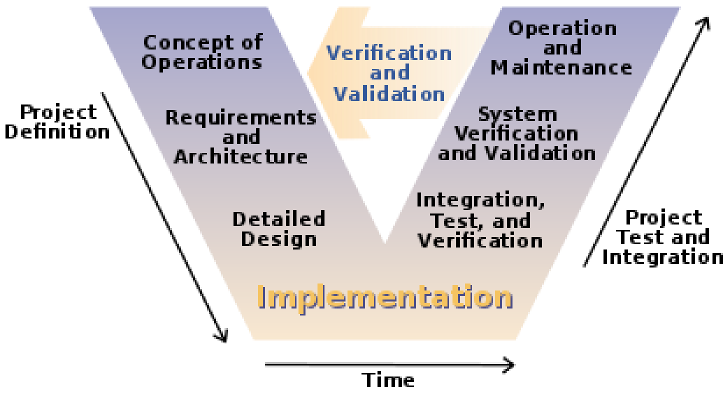

Figure 1. Systems Engineering Vee Model

The Vee Model is an iterative process looping through the verification steps multiple times. This iterative verification process starts with defining verification activities as part of defining system requirements, continues through integration and test of several implementation increments, and ends with the formal verification tests. See (Cloutier et al. 2019) for more information on the systems engineering Vee Model.

Problem Context

Requirements, verification plans and test scenarios are changing, and will then trigger a need for new test execution and analysis. A specific verification status is set on each requirement and verification plan test case (ok, not ok, not started), which will switch back to not started when a requirement, verification plan test case, or test scenario is changed.

Test execution of today in KDA has a limit regarding labor for manual operations. This inflict both stressful periods for the testers, and prevents full use of the test campaign timeframe. It is costly to have personnel work overtime, in addition to laws and regulations restricting working hours.

The test system of today produces more test results than the project organization can analyze, so the main bottleneck in the KDA test regime has proven to be the analysis of the test results. The analysis part is a tedious manual process, and insufficient analysis of system tests prior to the verification testing results in lack of error identification. Today's test system treats all tests equally by logging and storing test results. The test system issues no warning of any kind if the test has an unexpected result, other than compliance status to the defined acceptance criteria. Sub-system owners that have the responsibility of manually analyzing the test results are not automatically alerted of new test results, and often do not have time to look into it because of their own tight schedule.

For a given period, the JSM project performed 8-10 manual tests within normal working hours and 3-4 manual tests outside normal working hours. Three out of six available testers needed to be present to run two different test arenas. The test team arranged a shift regime from Monday to Thursday, with test periods from 8-12, 12-16 and 16-20. Friday and Saturday had only two shifts, and Sunday had none. On the other hand, an automated regime of test execution could run up to 200 tests per day. One test in one of the test arenas take about 30 minutes, while one test in the other test arena take roughly 10 minutes. A batch could have been set to run tests continuously for 24 hours, 7 days a week. On average one test needs one hour of manual analysis, which then results in 7-8 manual analysis per day, within normal working hours, performed by multiple analysts with different areas of expertise. For comparison, an automatic regime of test result analysis would quickly produce the test result analysis after each test. If one test fails, it could easily take several days to sort it out (error identification, corrective actions, re-test and re-analysis). Machine Learning could help with instantaneous error identification, speeding up the test failure process significantly. The manual test execution and test result analysis is very vulnerable to personnel getting sick, needing breaks (coffee, smoke, toilet, food and drink), needing sleep, getting bored, performing human errors, attending meetings, family matters, changing jobs, retiring, etc. The list of human weaknesses compared to machines is long.

Methods

To be able to find the best solution for making necessary improvements to the test system for product verification, exploring the problem is essential. Using a Systems Thinking approach will be a good way to achieve this understanding. Utilization of three methods to analyze the test system should be sufficient to gain the necessary insight. The CATWOE method exemplifies some different points of view for the test system from different stakeholders. See (Dharmalingam 2018) for more information on the CATWOE method. The systemigram method as proposed by Boardman helps understanding what affects the problem domain and how the solutions proposed affects the company. See (Boardman & Sauser 2008) for more information on the systemigram method. The causal loop method as described by Sterman shows how the test system dynamics reinforce and balance the forces in motion, and how the proposed changes affect these forces. See (Sterman 2000) for more information on the causal loop method.

(Dalkin et al. 2018) uses the CATWOE method to interact with stakeholders actively to overcome complexity related challenges. (Blair, Boardman, & Sauser 2007) uses the systemigram method to visualize the complexity of the System of Systems (SoS) in question systematically, and further using it as an effective platform for stakeholder dialog. (Neely & Walters 2016) uses causal loops to analyze a water supply system, and identifies leverage points to improve the structure of this water supply system for better sustainability.

By starting out defining the problem, before identifying the stakeholders and their needs, the development of the proposed test system used the Systems Engineering process.

By drawing a stakeholder interest map, defining value added processes and analyzing these using methods like CATWOE, systemigrams and causal loop diagrams, the development of the proposed test system used the Systems Thinking approach.

System Description

The system is complicated because of several interdependencies. The company must consider many interactions and effects when changing the test system.

The system is complex because of people involvement, which can lead to fault data (on purpose or not). Minimize people interaction through automation of some tasks could be done to make it less complex. An assumption is that less people interaction will reduce bottlenecks, manipulation and failure points/opportunities.

The test system consists of a structure, including different sub-systems, serving the necessary functionality to perform desired level of testing. The following sub-sections will describe the different sub-systems of the test system, as shown in Figure 2.

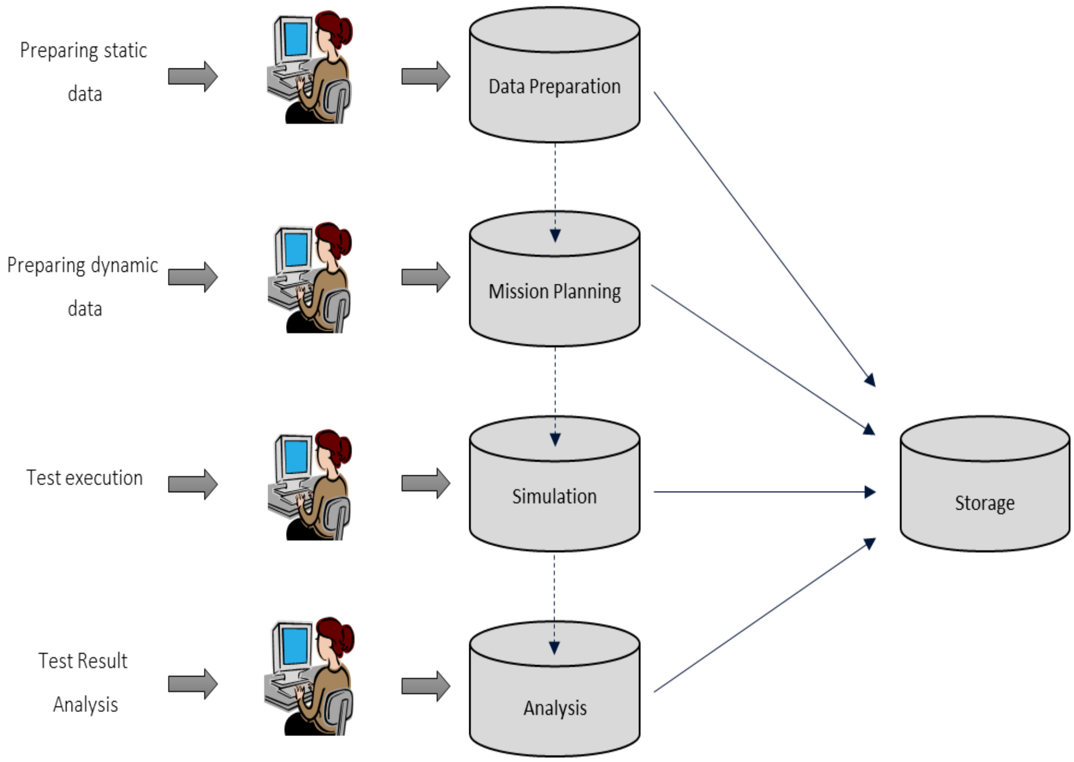

Figure 2. Current Test System

Data Preparation System:

A system that operators use to store intelligence data and adapt the data to relevant products. This could be map data, threat data and target data. People with specific domain knowledge will collect, extract and adjust data into different libraries for later retrieval in mission planning.

Mission Planning System:

A system that operators use to plan missions in advance to save time in the later mission execution phase. It is also an arena for testing and training of operational procedures. The mission planning system will have the same functionality in an off-board (ground facility) setting as on-board the launching platform.

Simulators. There are four different simulator environments serving different purposes in the test and verification process.

1) Host open loop test system is a simulator running in non-real-time, consisting of hardware (HW) and software (SW) to test the entire missile functionality up until missile launch. The main usage of this test arena is for initial testing.

2) Host closed loop test system is a simulator running in non-real-time, consisting of HW and SW to test the entire missile mission functionality. The main usage of this test arena is for initial testing and verification testing in need for batch simulations.

3) Target open loop test system is a simulator running in real-time, consisting of HW and SW to test the entire missile functionality up until missile launch as realistic as possible. The main usage of this test arena is for verification testing.

4) Target closed loop test system is a simulator running in real-time, consisting of HW and SW to test the entire missile mission functionality as realistic as possible. The main usage of this test arena is for verification testing and demonstrations.

To perform necessary manual operations, two testers are necessary in the last test arena during test execution, while one tester is enough for the other three test arenas.

Data Analysis System. Analyzers manually compare log file data from tests with acceptance criteria for the test cases in the verification plan.

Storage System. A storage system with enough space to store relevant data for later retrieval. The storage system will be a common server for test input and output data, while a separate product data management (PDM) system store the test documents.

Stakeholders

Stakeholder Identification

See Table 1 for identification of the stakeholders of the test system.

Table 1: Stakeholders and Their Interests

Stakeholders	Interests
Testers	The testers need to be able to use the test system to produce test results
System owners	The system owners need the test system to run specified scenarios and analyze the results according to their defined acceptance criteria
Project managers	The project managers need the test system to produce test reports fulfilling defined acceptance criteria
Company	The company need the test system to be able to verify products
Customers	The customers need the test system to verify the specified system requirements for the product to work as specified
Government	The government need the test system to contribute to product and project success, to make money from sales
Suppliers	The suppliers need the test system to require their hardware

Stakeholder Description

The testers are responsible for test execution in the different simulators, as well as maintaining and integrating the test structure. In addition, they report suspicious test results.

The system owners are responsible for the test description for their system, as well as the corresponding test report. They must provide specific test needs to the test team to ensure test execution. They also have the responsibility for analysis of test results concerning their system.

The project managers are responsible for the verification plan of their project, as well as the corresponding verification cross reference matrix. They are also responsible for agreement with the customers regarding necessary confidence in simulators used in test execution.

The company owns the test system, and the test system affects them based on the results and reputation established throughout different projects. They have the funding responsibility when it comes to upgrading the test system.

The customers own the products verified by the test system. A potential dysfunctional product directly affects the customers, if the testing has been insufficient. They must approve the verification plan and test descriptions before verification testing commence.

The government is the main shareholder in the company, and the test system affect them based on the test system impact on price and reputation. They also see to compliance with all laws and regulations.

The suppliers benefit from the test system need for original equipment, as well as later upgrades. They supply the company with HW for the test system and are responsible for providing necessary HW for the entire product life cycle.

Stakeholder Interest Map

The different sub-systems connect to the test system, as well as to other relevant sub-systems. Of the stakeholders, the entire test system influence the company, the government and the suppliers. Sub-systems of the test system influence the other stakeholders. Stakeholders also affect each other, forming the outer relationships of the stakeholder interest map. Figure 3 shows the stakeholder interest map.

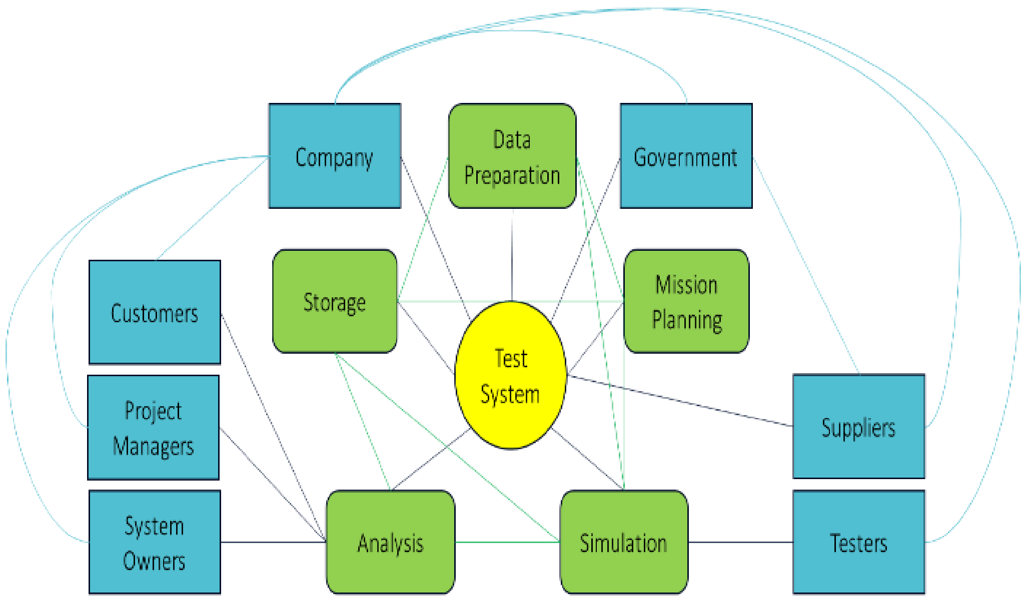

Figure 3. Stakeholder Interest Map

Proposed Changes

Value Added Processes

There are several value added processes foreseen to improve the company's test system of today. These will require work and introduce risk, but are estimated to be both necessary and to pay off in the future. The company will simply not be able to cope with future requirements for faster project execution and running more projects simultaneously using the test system of today. Figure 4 shows the proposed test system.

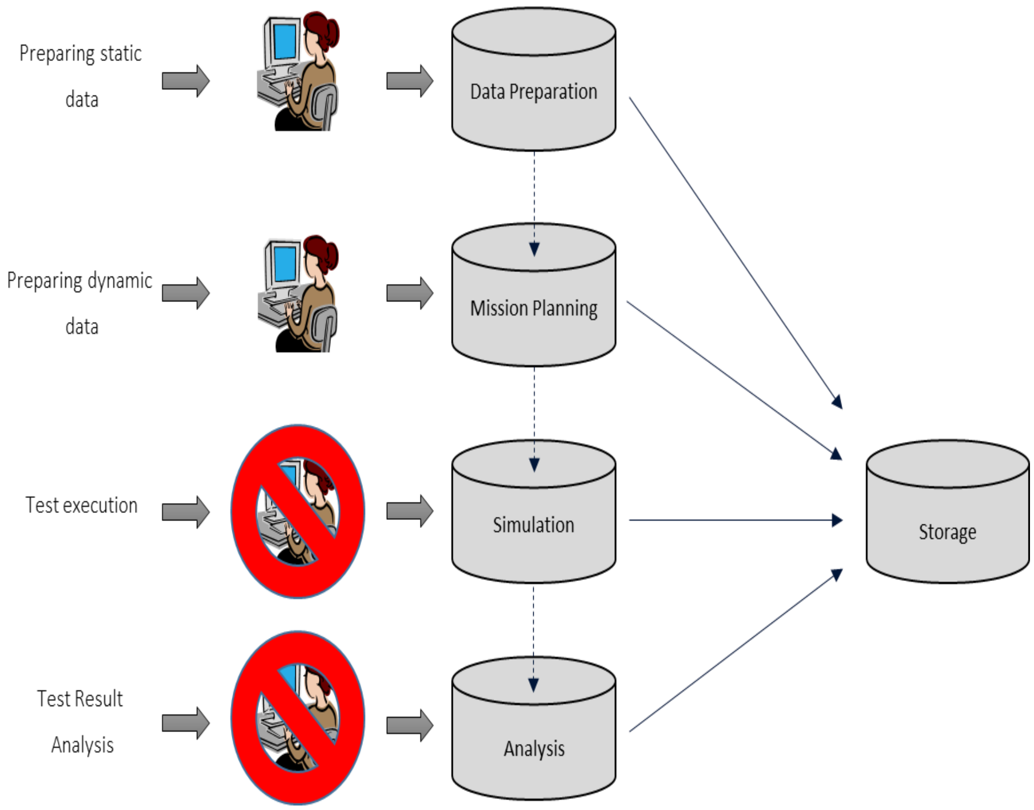

Figure 4. Proposed Test System

Automation of test execution and test result analysis will reduce the need for labor and save significant amount of time in each test campaign. In the future, the company should run the tests automatically by triggering test execution based on specified criteria. One person can optionally monitor the test execution from outside the lab. Automation of the test execution frees personnel from routine tasks to serve better purpose in other development areas. Proposed is a system for automatic analysis of available test results, preferably alerting operators about non-expected results. Analyzers make scripts in a tool like Matrix Laboratory (MATLAB) for the analysis system to fetch relevant data from the log files created after the tests. This functionality will require effort to implement, but will save significant time and money in future projects. See (Hahn & Valentine 2016) for more information on MATLAB.

Another improvement is the separate verification of core product and adaption product. The adaption product could consist of several special features not part of a standardized modular core product. Analysis on both core product and adaption product level will reduce the number of necessary tests in later projects, because the core product verification ensures a verified core product.

Combining test arena input/output over different systems, sub-systems and components will reuse data, preventing unnecessary documentation and testing.

Machine Learning will change the focus of everything to managing the exceptions, to cope with a large set of test results, only raising attention to problem areas. See (Alpaydin 2020) for more information on Machine Learning.

Obstacles

This value added processes will improve the test system for missile product verification, but some forces are preventing these changes from happening.

Lack of both domain knowledge and resources (people, time, money) are factors that the company must be prepared to deal with to be able to succeed in making these changes.

Lack of willingness to introduce risk for managers at different levels is an aspect that must be subject for discussion at company level to form a common policy in such matters.

Focus on project, not product, is a highly relevant topic. The company evaluates project managers based on project success, which changes the focus from product maturity to project milestones.

Overcoming these obstacles can be a much more difficult task than making the desired changes to the test system.

Considerations

The test system has to consider certain aspects in future development, restricting the degree of freedom.

Facilities like location and space are limiting factors for the test system. There is limited space in the test lab to set up multiples of different kinds of test stations. Type and number of test stations are restricting the possibilities for the test system. The test stations need to consist of HW capable of real time realistic testing. There must also be enough test stations to handle test demand within given periods, including redundancy in case of test station failures and scheduled/unscheduled maintenance.

The environment experienced by the testers is a factor, but will be removed or reduced with the proposition to remove the testers from the test arenas. The test stations produce noise that for a human being is undesirable to be working in for a longer period.

Security is yet another aspect that needs to be treated. The test stations must be able to handle classified data (up to NATO SECRET) and be able to treat it according to rules and regulations.

Safety is necessary to take seriously. The company must certify all testers for work in the test lab, and no tester should work in the test lab alone.

The company must consider all these aspects when deciding what proposed changes to the test system they want implemented, how, and to what level. Not considering these aspects might end up being very costly.

Analysis

CATWOE

A CATWOE analysis brings up the different stakeholders’ perceptions on a common platform, and provides a holistic understanding (Dharmalingam 2018). The CATWOE method is a novel and creative way of active engagement with stakeholders (Dalkin et al. 2018, p. 95). CATWOE is an acronym containing description of the following aspects: Customers, Actors, Transformation, Worldview, Owner and Environment. Figure 5 gives an overview of the CATWOE method (Dharmalingam 2018).

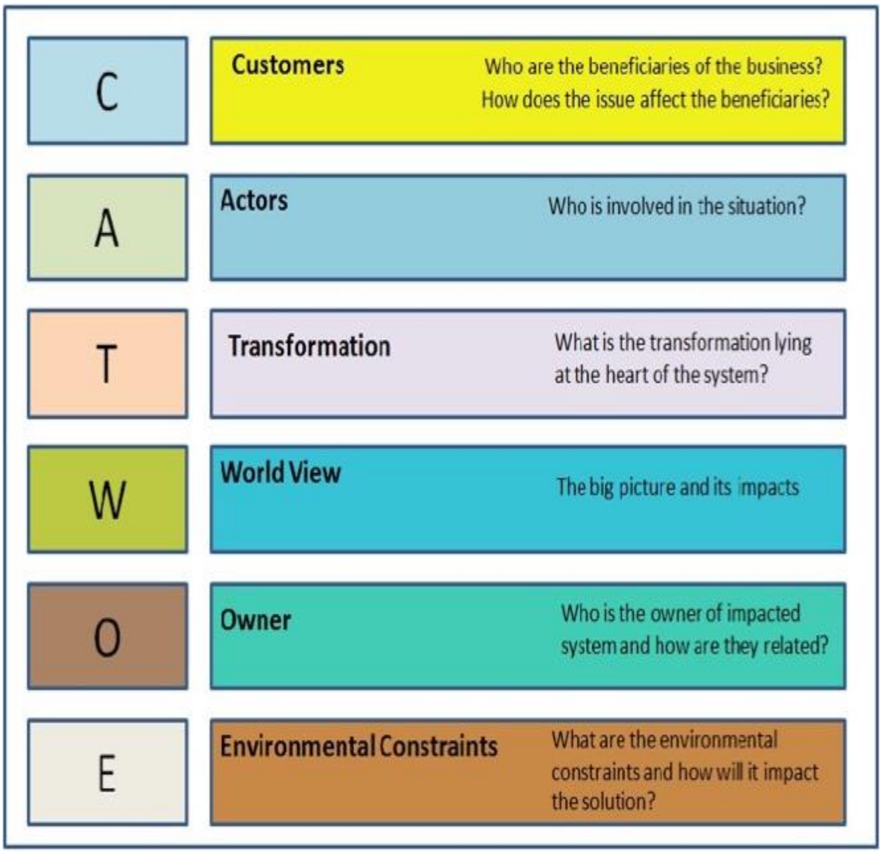

Figure 5. CATWOE

A CATWOE can be set-up for each stakeholder, which will have a somewhat different point of view. See Table 2 and Table 3 for examples of CATWOE for stakeholders of the test system.

Table 2: CATWOE Tester

Aspect	Description
Customers	Project Manager
Actors	SW developers, scenario data providers, testers and analyzers
Transformation	Provides test results based on scenario data and analysis of test results based on log files
Worldview	Simulation of test scenario
Owner	Project
Environment	Test stations (HW, SW, noise)

Table 3: CATWOE Project Manager

Aspect	Description
Customers	Norwegian Armed Forces

Aspect	Description
Actors	System owners and testers
Transformation	Provides test results based on requirements and verification plans
Worldview	Verification of product
Owner	Company
Environment	Test lab (number and type of test stations)

The CATWOE method exemplifies that the different stakeholders can have different aspects to the system in question. This can again result in different abstraction levels in their needs, even though the proposed solution may fit both. This research used the CATWOE method to communicate and understand the stakeholders' perspectives, but it is necessary to look deeper into the interactions of the test system to understand it better. The systemigram method helps achieve this deeper understanding.

Systemigram

Systemigram is a novel medium for capturing strategic intent, in a way that prepares the ground for consensus building among diverse communicants (Boardman & Sauser 2008, p. 111).The systemigram gives an overview of the test system elements and their interactions, which helps in understanding how changes will affect the system. Dependent on where the change is applied, the change could be isolated or causing a chain reaction. See Figure 6 for the systemigram of the current test system, and Figure 7 for the systemigram of the proposed test system.

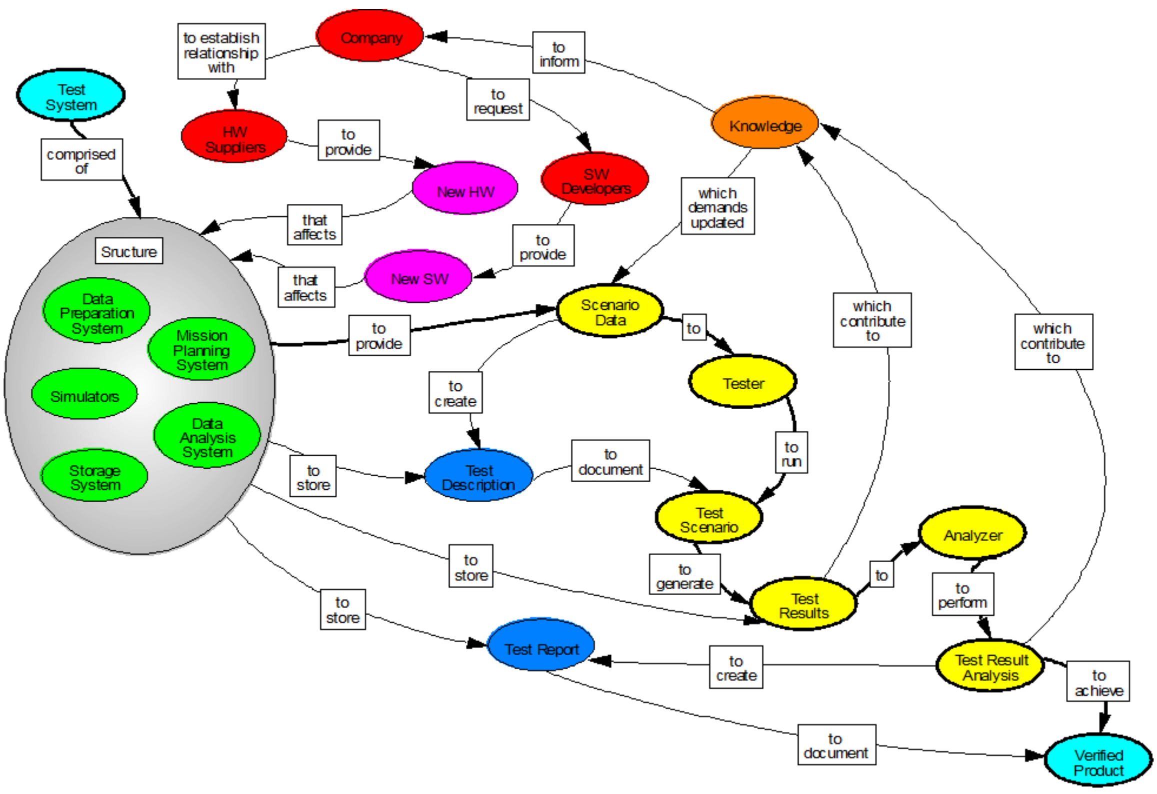

Figure 6. Systemigram Current Test System

The mainstay gives the main steps of the story for the system in question to reach the end goal. This is the essence and are the necessary steps for the test system to reach the goal of a verified product.

For the current test system, the mainstay reads *Test System comprised of Structure (Data Preparation System, Mission Planning System, Simulators, Data Analysis System and Storage System) to provide Scenario Data to Tester to run Test Scenario to generate Test Results to Analyzer to perform Test Result Analysis to achieve Verified Product.*

Other loops in the systemigram support and affect the mainstay. 1) *Test Results, which contribute to Knowledge, which demand updated Scenario Data.* 2) *Test Results Analysis, which contribute to Knowledge, which demand updated Scenario Data.* 3) *Test Results, which contribute to Knowledge, to inform Company to request SW Developers to provide New SW that affects Structure.* 4) *Test Results Analysis, which contribute to Knowledge, to inform Company to request SW Developers to provide New SW that affects Structure.* 5) *Test Results, which contribute to Knowledge, to inform Company to establish relationship with HW Suppliers to provide New HW that affects Structure.* 6) *Test Results Analysis, which contribute to Knowledge, to inform Company to establish relationship with HW Suppliers to provide New HW that affects Structure.* 7) *Scenario Data to create Test Description to document Test Scenario.* 8) *Test Results Analysis to create Test Report to document Verified Product.* 9) *Structure to store Test Description.* 10) *Structure to store Test Results* 11) *Structure to store Test Report.* These steps are supporting or optional operations of the test system. Storage of test results is an example of supporting functionality, while a necessary change in SW is an example of optional functionality.

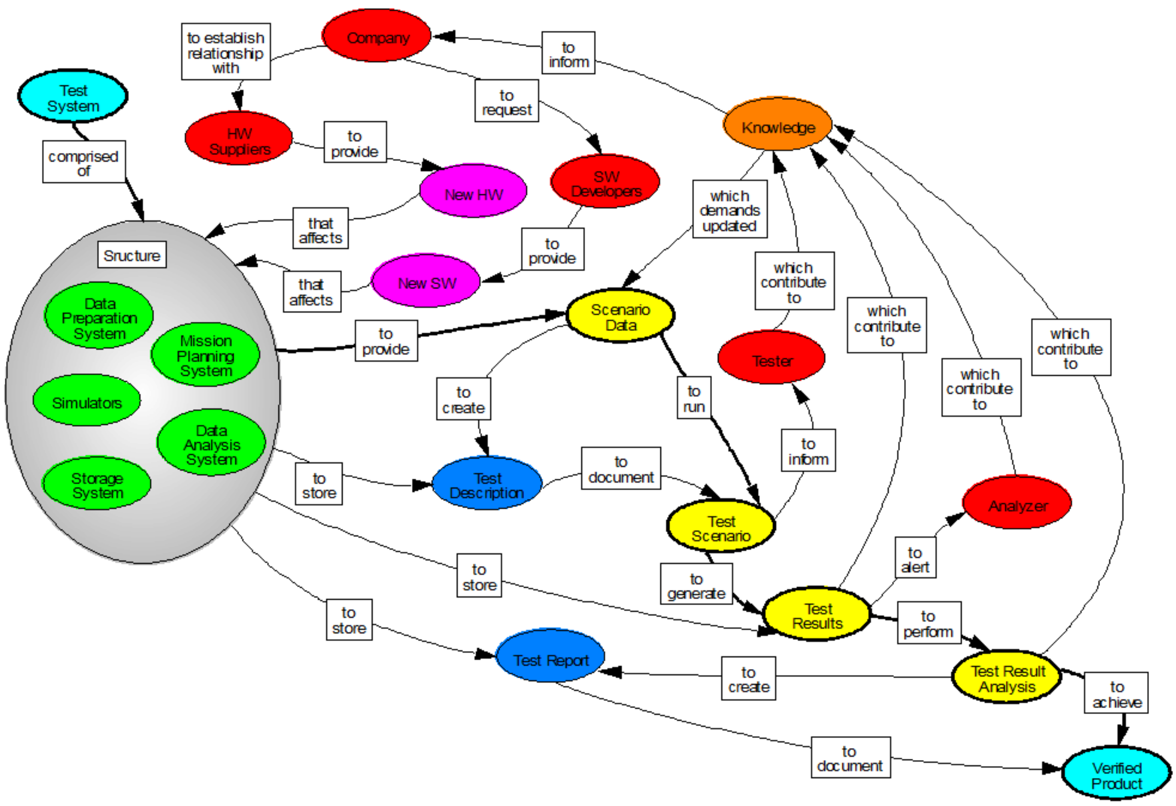

Figure 7. Systemigram Proposed Test System

For the proposed test system, the mainstay reads *Test System comprised of Structure (Data Preparation System, Mission Planning System, Simulators, Data Analysis System, and Storage System) to provide Scenario Data to run Test Scenario to generate Test Results to perform Test Result Analysis to achieve Verified Produc*t.

The proposed test system has additional optional loops in the systemigram compared to the current test system. 1) *Test Scenario to inform Tester which contribute to Knowledge which demand updated Scenario Data.* 2) *Test Scenario to inform Tester, which contribute to Knowledge, to inform Company to request SW Developers to provide New SW that affects Structure.* 3) *Test Scenario to* *inform Tester, which contribute to Knowledge, to inform Company to establish relationship with HW Suppliers to provide New HW that affects Structure.* 4) *Test Results to alert Analyzer, which contribute to Knowledge, which demand updated Scenario Data.* 5) *Test Results to alert Analyzer, which contribute to Knowledge, to inform Company to request SW Developers to provide New SW that affects Structure.* 6) *Test Results to alert Analyzer, which contribute to Knowledge, to inform Company to establish relationship with HW Suppliers to provide New HW that affects Structure.*

The systemigram method has helped seeing how the test system works and how changes take effect. A fully automated test execution process would require the scenario data to be unambiguous, making a computer able to read it without any need for interpretation. An operator must describe the scenario data in a way that a computer can understand it in the same way as a human operator. The need for testers would be lower, reducing labor hours and at the same time broadening the time of operation to be continuous. A fully automated analysis process of the test results would require a significant effort to establish the necessary analysis scripts. These scripts must contain sufficient checks according to defined acceptance criteria, and be robust to continuous testing. The need for analyzers would be lower, reducing labor hours and at the same time broadening the time of operation to continuous. A Machine Learning process would require a significant effort to establish the Machine Learning regime, as well as later follow-up/maintenance to make desired adjustments to the algorithms. Machine Learning will reduce the risk of late discovery of errors, by alerting the analyzer about suspicious test results during the entire test campaign.

The systemigram helps to investigate the effect of removing the manual operations of test execution and test result analysis from the critical line, and further explore how the dynamics of the test system are changed. The test personnel are going from necessary and time-consuming routine tasks to more of an observer role. The test personnel can then use more of their time on new development and other optimization type of work, which in most cases are rated as more interesting type of work for engineers.

Causal Loop Diagrams

In the domain of system dynamics, causal loop diagrams (CLDs) qualitatively present the dynamic influences between factors thought to influence a particular system behavior (Neely & Walters 2016). Usage of the causal loop method take the next step in exploring the mechanisms of the test system, by looking at the forces in motion, how they reinforce and balance.

There will be more test data available as more test results are ready, which again will increase the test maturity. On the other hand, the test structure will give restrictions in number of test runs during a defined period, which again depend on available simulators at any given time and limit the amount of test results. For the current test system, the tester will also be a limiting factor, but the proposed test system removes this factor. See Figure 8 and Figure 9 for an illustration of the test execution causal loops for the current and proposed test system.

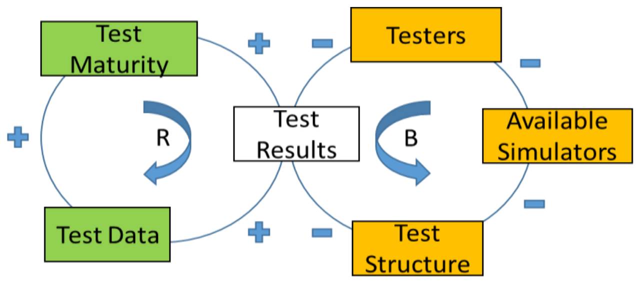

Figure 8. Causal Loop Test Execution Current Test System

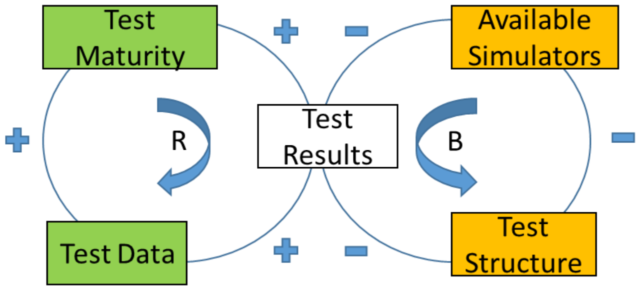

Figure 9. Causal Loop Test Execution Proposed Test System

The reinforcing factor R is larger than the balancing factor B in the current test system, since the test maturity is continuously increasing. The removal of the tester in the proposed test system will at least triple the test executions for a given period, based on continuous testing compared to testing during normal working hours. The balancing loop will then be three times weaker (B/3), which will increase the test maturity level by a factor of three.

There will be more test reports ready as more test analysis are ready, which again will increase the verification level. On the other hand, the analysis structure will dictate the extent of automatic analysis, which again affect necessary time for test analysis. If the test system does not reach the goal of full automation of the test results analysis process, this will require a significant amount of time for manual test results analysis. For the current test system, the analyzer will be a limiting factor, but the proposed test system removes this factor. See Figure 10 and Figure 11 for an illustration of the test analysis causal loops for the current and proposed test system.

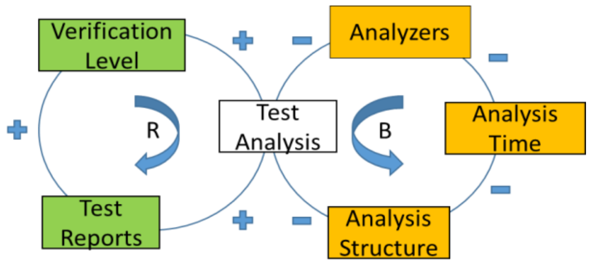

Figure 10. Causal Loop Test Analysis Current Test System

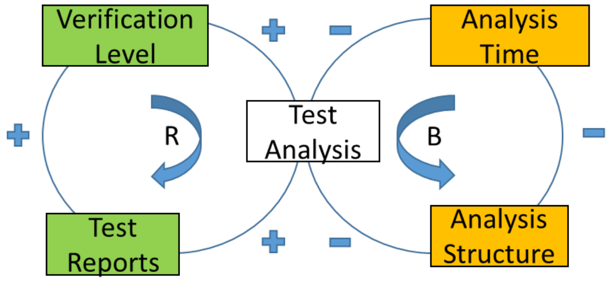

Figure 11. Causal Loop Test Analysis Proposed Test System

The reinforcing factor R is larger than the balancing factor B in the current test system, since the verification level is continuously increasing. The number of tests that need analysis will control the effect of removing the analyzer. The time to do a manual test result analysis can vary from a couple of minutes to multiple hours, based on experience from previous test periods. A large project will then save significant time in the test analysis process.

Introduction of Machine Learning could increase probability of finding errors early in the test campaign instead of late in the verification process. This will reintroduce the analyzer role, but the benefit of early error detection is much higher than the negative effect of increase in time for analysis. Error detection in the verification process will cost 50% more, based on work outside normal working hours.

Previous steps in the Vee Model often delay the final product verification so much that there is no room for errors to be able to finish in time, only depending on work within normal working hours. Personnel salary increase by 50% outside normal working hours, and even increase by 100% during weekends.

See Figure 12 for an illustration of the Machine Learning causal loop for the proposed test system.

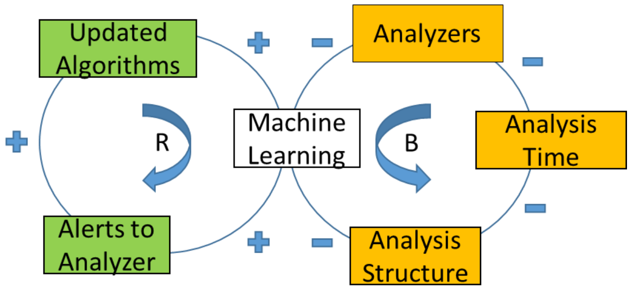

Figure 12. Causal Loop Machine Learning Proposed Test System

The reinforcing factor is larger than the balancing factor for the Machine Learning, since the algorithms will be continuously evolving. Developers will update the Machine Learning algorithms as analyzers evaluate alerts to be valid or not.

The causal loop method helps seeing how the forces in motion are reinforced and balanced. This gives a clear indication of the potential benefits of the proposed test system compared to the current test system. Removing manual operations will reduce the balancing forces in motion for both test execution and test result analysis, which again leads to increased test maturity and verification level. An improvement factor of at least three applies due to the inherent capacity of machines to work continuously, while people normally just work for 8 out of 24 hours. Taking into consideration all benefits of machines over humans and looking at numbers retriever from earlier test periods with manual operations, a more realistic improvement factor of fifteen seems reasonable. Previous work on automated testing to improve efficiency can support the problem defined in this paper. (Enoiu et al. 2017, p. 1) estimates a test coverage improvement of roughly 90% within the same period, going from manual to automatic testing.

Leverage Points

Leverage points of a system are elements that can make a significant impact, positive or negative. The test system has two identified leverage points.

The first leverage point of the test system is the test structure. The test structure can be a bottleneck for a rapid verification process, but can change to provide necessary capacity. Removing the testers from the simulators in the test structure for the proposed test system would have a significant positive impact.

The second leverage point of the test system is the analysis structure. The analysis structure can be a limiting factor for a rapid verification process, but can change to provide an accelerated verification process. Removing the analyzers from the data analysis system in the test structure for the proposed test system would have a significant positive impact. The company could further improve the test system's analysis structure by introducing Machine Learning. This will reintroduce the analyzer role, but to a less extent than before. Machine Learning can help significantly in coping with a reasonable amount of test data and relevant issues, by alerting about suspicious test results. These alerts can help identify errors early in the test campaign.

The proposed test system exploit all these leverage points, significantly increasing the capability and capacity.

Conclusion

The company's test system is vital for product verification. Today it requires manual operations, which has proven to be bottlenecks in the test execution and test results analysis processes. Usage of a Systems Thinking approach have explored this problem and helped see the effects of proposed changes. Methods used to analyze the problem are CATWOE, systemigram and causal loops. The CATWOE method has helped communicating with and understanding stakeholders of the test system. The systemigram method have visualized how elements of the test system are connected, how they affect each other and how the proposed changes impact. The causal loops method have stated how the forces in motion balance each other and how the proposed changes affect this balance.

To exemplify that different stakeholders take on somewhat different abstraction levels on different aspects of the system, even though they have the same goal, the CATWOE method is used.

The analysis by use of systemigram has visualized the effect of removing personnel from the routine tasks of test execution and test result analysis, having personnel take more of an observer role. This will result in personnel being able to use their time on other more interesting tasks like new development and optimization. Further analysis by use of causal loops have shown that the reinforcing forces in motion to improve test maturity and verification level increase as personnel are removed from the balancing forces in motion. Engineering judgements based on this analysis estimate an improvement factor of three to fifteen, based on how much faster and more robust machines perform than humans.

The company should upgrade its test system to cope with future requirements for faster project execution and running multiple projects in parallel. The test system should have an automated test execution and test results analysis process, not requiring tedious manual operations. The automated analysis process should further introduce Machine Learning to change the focus of everything to managing the exceptions, to alert only suspicious test results subject for further manual analysis.

The type and level of proposed changes to the current test system are not in conflict with the defined considerations, but are with the defined obstacles. The company management need convincing to form policies for the proposed changes to the test system to become realistic.

A more extensive literature study to review related work on test automation would be beneficial to take advantage of previous shared knowledge. A dedicated comparison type of analysis between manually performed tests and test result analysis versus automatically performed tests and test result analysis would strengthen the paper, and will be a natural part of future work.

References

Alpaydin, E. 2020, *Introduction to Machine Learning*, 4th edn, The MIT Press, Massachusetts.

BKCASE Editorial Board. 2019. *The Guide to the Systems Engineering Body of Knowledge (SEBoK)*, v.2.1, R.J. Cloutier (Editor in Chief). Hoboken, NJ: The Trustees of the Stevens Institute of Technology. Accessed 01 April 2020. <www.sebokwiki.org>. BKCASE is managed and maintained by the Stevens Institute of Technology Systems Engineering Re-search Center, the International Council on Systems Engineering, and the Institute of Electrical and Electronics Engineers Computer Society.

Blair, C., Boardman, J., and Sauser, B. 2007, *Communications Strategic Intent With Systemigrams: Application to the Network-Enabled Challenge*, Systems Engineering. 10.309-322.10.1002/sys.20079.

Boardman, J., and Sauser, B. 2008, *Systems Thinking: Coping with 21st Century Problems*. Boca Raton, FL: Taylor & Francis / CRC Press.

Dalkin, S., Lhussier, M., Williams, L., Burton, C. R., and Rycroft-Malone, J. 2018, *Exploring the use of Soft Systems Methodology with realist approaches: A novel way to map programme complexity and develop and refine programme theory*. Evaluation, 24(1), 84–97. <https://doi.org/10.1177/1356389017749036>.

Dharmalingam, N. 2018, *10 Most Popular Business Analysis Techniques,* viewed 29 July 2019 <https://www.whizlabs.com/blog/best-business-analysis-techniques/>.

Enoiu, E., Sundmark, D., Causevic, A., and Pettersson, P. 2017, *A Comparative Study of Manual and Automated Testing for Industrial Control Software*. 412-417.10.1109/ICST.2017.44. Conference: 2017 IEEE International Conference on Software Testing, Verification and Validation.

Hahn, B.H., and Valentine, D.T. 2019, *Essential MATLAB for Engineers and Scientists*, 7th edn, Academic Press.

Neely, K., Walters, J.P. 2016, *Using Causal Loop Diagramming to Explore the Drivers of the Sustained Functionality of Rural Water Services in Timor-Leste,* Sustainability 2016, 8, 57. <https://doi.org/10.3390/su8010057>.

Osborne, L., Brummond, J., Hart, R., Zarean, M. and Conger, S. 2005, *Clarus Concept of Operations*, Doctoral Dissertation, Publication No. FHWA-JPO-05-072, Federal Highway Administration (FHWA). < https://rosap.ntl.bts.gov/view/dot/3710>.

Project Management Institute 2018, *A guide to the project management body of knowledge (PMBOK guide)*, 6th edn, Newtown Square, PA, United States.

Sterman, J.D. 2000, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, Boston.

Biography

Rune Andre Haugen was in service with the Royal Norwegian Air Force (RNoAF) from 1997 to 2003, including graduation from the RNoAF Officer Candidate School in Stavern (1999) and the RNoAF Academy in Trondheim (2001). He holds both a Bachelor degree (2006) and a Master degree (2013) in Systems Engineering from USN. He has worked as a design engineer at FMC Kongsberg Subsea from 2006 to 2008 (3D modelling), and as a system engineer in KDA since 2008 (system design and test).

Mo Mansouri is an Assistant Professor in School of Systems and Enterprises at Stevens Institute of Technology. His research is focused on development of computational and quantitative governance frameworks for complex adaptive socio-technical environments as well as network industries. He holds a Bachelor degree from Sharif University of Technology, a Master degree from the University of Tehran, both in Industrial Engineering, and a Doctor of Science in Engineering Management from The George Washington University.